Economic Barriers to the Deployment of Existing Privacy Technologies (Position Paper)

Joan Feigenbaum[1], Michael J. Freedman[2], Tomas Sander[3], Adam Shostack[4]

[1] Yale University (joan.feigenbaum@yale.edu)
[2] MIT Lab for Computer Science (mfreed@lcs.mit.edu)
[3] InterTrust STAR Lab (sander@intertrust.com)
[4] Zero-Knowledge Labs (adam@zeroknowledge.com)

Internet-based commerce provides great opportunities for merchants, consumers, and business affiliates, but it may seriously threaten users' privacy. Some of the paths to loss of privacy are quite familiar (*e.g.*, mining of credit-card data), but some are new or much more serious than they were in earlier regimes. We present two points about the economics of electronic commerce: There are economic barriers to the adoption of privacy-enabling technology; furthermore, it is unclear that most of the privacy-enabling technology studied by the cryptology R&D community would address users' actual privacy concerns even if it were widely adopted. More details can be found in [2], which focuses on digital rights management (DRM) systems.

Twenty-five years of cryptographic research has yielded a vast array of privacy-enabling technologies that support many types of two-party and multi-party interactions. Thus, cryptographic researchers might wish to believe that user privacy in e-commerce and content distribution is a solved problem. You pay for content or services with anonymous electronic cash. You connect to content or service providers via an anonymizing mixnet. You authenticate yourself with anonymous credential schemes or zero-knowledge identification protocols. You download content via private information retrieval or oblivious transfer. You use secure function evaluation when interacting with services that require some information.

Despite the fact that many of the impressive techniques in the cryptographic research literature have been extensively and rigorously analyzed, and some have even been commercially developed, few are in widespread use. It is our thesis that there are straightforward economic and business reasons for this apparent contradiction.

The major constituencies involved in a privacy-enabling protocol or system must be willing to sacrifice the information that could be collected about the other parties or their inputs. In the absence of legal requirements – that are generally understood, technologically feasible, and consistently enforced – use of such protocols and systems must be voluntary and bilateral. However, in e-commerce transactions, these constituencies have conflicting interests and asymmetric power. Why should a powerful content/service provider wanting to learn information about his users agree to run a protocol that deprives him of this very information? Industry is likely to the follow the "Know your customer" mantra.

Many of the problems facing privacy-technology adoption can be framed in microeconomic terms, *e.g.*, network externalities, asymmetric information, and moral hazard. See Anderson [1] and Shapiro and Varian [3] for similar arguments about related domains.

It is easy to see that many privacy technologies obey Metcalfe's law and therefore exhibit network externalities – their marginal value to a user increases with their expected number of users. Anonymous file-sharing systems will become truly beneficial to users only when a large array of content can be readily, easily accessed. Anonymous email is unidirectional (and therefore less useful) unless both parties use the anonymizing network. The anonymity offered by such a network is bounded by the number of users. Similarly, electronic cash will only become useful if many merchants will accept it. We may infer from this that DRM and other e-commerce systems are unlikely to push the acceptance of cryptographic ecash but rather will continue with existing payment methods such as credit cards.

Several other features of network economics are of particular importance. Technology often has high fixed cost and low marginal costs, and switching costs for infrastructural technologies are also quite large, leading to lock-in. Assuming that corporate entities make decisions motivated primarily by profit (and that a good reputation for respecting customers' privacy has a measurable positive impact on profitability), these entities should only switch infrastructural technologies if the expected net present value of the benefits of switching is greater than its costs. Experience shows that this makes infrastructural switching rare, slow, and painful.

Often, part of what makes a business an "Internet business" is that it can use pre-existing Internet infrastructure to get a cost advantage over its competitors. If privacy technologies require widespread infrastructure redesign, they vitiate this principle of Internet business success, and content/service providers probably will not adopt them. If ubiquitous onion routing requires changing network protocols and routers, and the only benefit is consumer privacy, we had better not have to wait for onion routing to be in place in order to be able to buy and read e-books in private!

An asymmetry of information between entities in an e-commerce system makes privacy more difficult to achieve. Moral hazard arises from the principal-agent problem, in which the principal (*i.e.*, consumer) cannot observe the effort of the agent (*i.e.*, content/service provider) and thus has to incentivize the agent using something other than a payment per unit of effort. The hazard arises when the results of the agent's effort (*i.e.*, the "amount" of privacy) cannot be measured accurately, and thus the agent is tempted to slack off. The obvious conclusion of this economic argument is that providers will be tempted not to provide privacy because consumers cannot really measure their "units of privacy" and make educated demands.

Finally, note that there are legitimate reasons for businesses to collect data, including customer retention, statistics, risk management, customization, and billing. For instance, network operations can (and perhaps should) collect usage data for traffic-modeling and provisioning purposes. Lack of good Internet traffic models is a big problem, and Internet-traffic modeling is a very active area of research; it requires the collection of usage data.

Our abstractions don't model our reality

The cryptographic research community models interactions and protocols in terms of very distinct entities and information. For instance, the traditional communication confidentiality model is that Alice wants to communicate with her friend Bob without adversaries Eve and Lucifer being able to read her message. We may abstract general privacy-enhancing protocols by saying that users try to hide information by performing computations in some trusted private environment (the "trusted computing base," or TCB) and then using the results of these computations to communicate with the outside world.

DRM enters the picture, for example, because users want to obtain mass-market content online and commercial distributors want to sell it to them. Many users are unconcerned about the commercial distributors knowing the details of the purchases in which they participate directly and using this knowledge for straightforward business purposes (such as order fulfillment and billing), but many are concerned about how such knowledge could be misused or shared. This problem is further complicated by the fact that "misuse" is ill-defined; pairs of parties that have some interests in common also have some conflicting interests. Alliances, partnerships, and commercial relationships are constantly in flux and will remain so. Not only are users battling from the low ground, but it is difficult for them even to identity the enemy from whom they should hide all of their private data. In a network where businesses bundle valuable content and personalization services, and users want anytime anywhere access from any number of devices, who is Alice, who is Bob, who is the enemy, and what is the TCB? Cryptographic research cannot answer these questions.

Cryptographic-protocol specifications assume that one knows exactly what is "legitimate use" of data and what is "misuse", and they assume that there is a single, well-defined relationship between Alice and Bob. We suggest that this model is inadequate for electronic commerce, where the clean dichotomies of good guy vs. bad guy, trusted vs. untrusted, and private vs. public do not exist.

References

1. R. Anderson, "Why information security is hard - an economic perspective," Available at `http://www.cl.cam.ac.uk/ftp/users/rja14/econ.pdf`.
2. J. Feigenbaum, M. Freedman, T. Sander, and A. Shostack, "Privacy Engineering for Digital Rights Management Systems," to appear in Proceedings of the 2001 ACM Workshop on Security and Privacy in Digital Rights Management. Available in preprint form at `http://www.cs.yale.edu/homes/jf/FFSS.ps`.
3. C. Shapiro and H. Varian, Information Rules: A Strategic Guide to the Network Economy, Harvard Business School Press, Boston, 1999.